北京化工大学



题　目　基于 ResNet 网络的花卉种类识别研究

专　业　信息与通信工程

班　级　信研 2308 班

姓　名　张迎松

学　号　2023200852

2023 年 12 月 31 日

基于 ResNet 网络的花卉种类识别研究

摘 要

近年来，随着深度学习技术的迅速发展，基于卷积神经网络（CNN）的图像识别在各个领域取得了显著的成就。在植物学领域，花卉种类的识别是一个重要的研究方向，对于生态学、农业以及环境监测等诸多方面具有重要意义。本研究旨在探索并优化 ResNet（深度残差网络）在花卉种类识别任务中的应用。首先，文章对 ResNet 网络的结构进行了深入分析，理解了其引入残差学习的机制，以及如何有效地应对深层网络训练中的梯度消失和爆炸问题。通过在花卉图像数据集上的实验，验证了 ResNet 在处理复杂多类别花卉图像识别任务上的卓越性能。在数据预处理阶段，文章采用了数据增强技术，包括裁剪和翻转等，以扩充训练数据集，提高模型的泛化能力，同时对花卉图像进行了标准化处理，以适应 ResNet 网络对输入数据的要求。实验结果表明，相较于传统的神经网络模型，采用 ResNet 网络的花卉种类识别模型在准确率和收敛速度上均有显著提升。此外，通过深度分析模型在不同花卉类别上的表现，发现 ResNet 网络在处理具有层次结构和复杂形态的花卉图像时表现更为出色。本文提出的模型不仅在整体上取得了优异的性能，还在对特定花卉类别的识别上具备较高的准确性。在进一步的研究中，我们考虑通过迁移学习进一步提升模型的泛化能力，特别是在面对小样本花卉数据集时。同时将探索模型的实时性能，以适应现实场景中对花卉种类快速准确识别的需求。

本研究通过对 ResNet 网络在花卉种类识别任务中的优势和应用进行全面而深入的分析，为深度学习在植物学领域的应用提供了有益的参考和借鉴。研究成果不仅对花卉识别技术的改进具有一定的理论价值，同时也在实际应用中具备广泛的推广潜力。

关键词：花卉识别，ResNet 网络，迁移学习

RESEARCH ON RESERVATION TYPE IDENTIFICATION BASED ON RESNET NETWORK

ABSTRACT

In recent years, with the rapid development of deep learning technology, image recognition based on convolutional neural networks (CNN) has achieved remarkable achievements in various fields. In the field of botany, flower species identification is an important research direction and is of great significance in ecology, agriculture, and environmental monitoring. This research aims to explore and optimize the application of ResNet (deep residual network) in flower type recognition tasks. First, the article conducts an in-depth analysis of the structure of the ResNet network, understands its mechanism for introducing residual learning, and how to effectively deal with the vanishing and exploding gradient problems in deep network training. Through preliminary experiments on a large-scale flower image data set, the excellent performance of ResNet in handling complex multi-category flower image recognition tasks was verified. In the data preprocessing stage, the article uses data enhancement techniques, including cropping and flipping, to expand the training data set and improve the generalization ability of the model. At the same time, the flower images are standardized to adapt to the requirements of the ResNet network for input data. Experimental results show that compared with the traditional CNN model, the flower type recognition model using ResNet has significantly improved accuracy and convergence speed. In addition, through in-depth analysis of the model's performance on different flower categories, it was found that the ResNet network performed better when processing flower images with hierarchical structures and complex shapes. The model proposed in this article not only achieves excellent performance overall, but also has high accuracy in identifying

specific flower categories. In further research, we consider further improving the generalization ability of the model through transfer learning, especially when facing small sample flower data sets. At the same time, the real-time performance of the model will be explored to adapt to the need for rapid and accurate identification of flower types in real scenes.

This study provides a useful reference and reference for the application of deep learning in the field of botany by conducting a comprehensive and in-depth analysis of the advantages and applications of ResNet network in flower species recognition tasks. The research results not only have certain theoretical value for the improvement of flower identification technology, but also have extensive potential for promotion in practical applications.

KEY WORDS: Flower Recognition, Resnet Network, Transfer Learning

目录

前 言

随着社会的进步和生活质量的提高，人们对花卉的的关注逐渐升温，花卉种类的识别对于植物学家、园艺师以及环境保护者具有重要价值。然而，由于花卉种类繁多、形态各异，传统的人工识别方法在效率和准确性上面临挑战。而随着深度学习技术的崛起，基于卷积神经网络（CNN）的图像识别在多个领域取得了显著的成功。虽然卷积神经网络在图像识别中表现出色，但是在处理深层网络中存在的梯度消失和梯度爆炸问题时，传统的网络结构逐渐显得力不从心。ResNet（Residual Network）由其引入的残差学习框架而在 ImageNet 竞赛中大放异彩，其独特的残差块设计有效地缓解了梯度问题，使得网络更深时反而更容易训练。因此，将 ResNet 应用于花卉种类识别，有望提高识别准确性并加速训练过程[1]。

本研究的主要目的在于探索基于 ResNet 网络的花卉种类识别方法，通过深度学习技术提高花卉图像的自动化分类准确性。通过借鉴 ResNet 网络的残差学习机制，本研究旨在建立一个稳健、高效的花卉识别模型，为植物学研究、生态监测和园艺设计等领域提供有力的支持。本研究将主要关注以下几个方面：

花卉数据集的构建：采用大规模、多样性的花卉图像数据集，以确保模型具有良好的泛化能力。

ResNet 网络的深入研究：通过对 ResNet 网络结构的详细分析，深入理解其残差学习机制，并针对花卉种类识别任务进行优化。

模型训练与优化：利用构建的花卉数据集对 ResNet 网络进行训练，并通过调整超参数、采用迁移学习等方法对模型进行优化，以提高识别准确性。

通过本研究，预期将为花卉种类识别领域的深度学习应用提供新的思路和方法。同时，本研究还有望为其他图像识别任务提供借鉴和参考，推动深度学习在植物学、农业和环境科学等领域的应用。

第 1 章 绪论

1.1 研究背景与意义

1.1.1 研究背景

花卉种类识别作为计算机视觉领域中的一个重要任务，近年来受益于深度学习技术的发展，取得了显著的进展。花卉识别在农业、生态学、园艺学等领域有着广泛的应用，例如在植物学研究、城市绿化规划和花卉市场管理等方面发挥着关键作用。然而，由于花卉的种类繁多、形态各异，传统的图像识别方法在解决这一问题上显得力不从心，而深度学习技术的兴起为图像识别领域带来了新的机遇。卷积神经网络（CNN）作为深度学习的代表，成功地解决了传统方法中需要手工设计特征的问题，而 ResNet 网络作为一种深度残差网络，在解决深度网络训练中的梯度问题上具有独特的优势。因此，将 ResNet 网络应用于花卉种类识别，成为提高模型性能和泛化能力的一个有前景的方向。传统的花卉种类识别方法通常依赖于手动设计的特征提取器，这些特征提取器对花卉图像中的复杂纹理、颜色和形状往往缺乏有效的刻画。因此，在面对大规模、高维度的花卉图像数据时，传统方法的性能容易受到限制。此外，传统方法通常需要大量的领域专业知识和经验，限制了其在实际应用中的普适性[2]。

深度学习的兴起为花卉种类识别带来了新的希望，深度学习模型能够通过端到端的学习，自动地从大量的数据中学习到高层次的抽象特征，无需手动设计特征提取器。这种自动学习的特性使得深度学习模型在处理复杂多变的花卉图像时表现更为出色。

1.1.2 研究意义

深度学习模型，特别是 ResNet 网络，能够自动学习图像中的抽象特征，从而提高花卉种类识别的准确性。花卉图像可能具有复杂的背景、多样的角度和光照条件，ResNet 网络引入的残差学习机制有助于处理这些复杂场景，提高模型的鲁棒性。深度残差网络的设计有效解决了深层网络训练中的梯度消失和爆炸问题，使得模型更易于训练，收敛速度更快，从而提高了训练效率。花卉种类识别技术的提升将有助于加深对植物分类学和植物生态学的认识，推动相关领域的研究进展。花卉种类识别技术在农业和园艺中有广泛的应用，通过精准地识别花卉种类，可以帮助农民和园艺爱好者更好地管理植物，提高生产效率和观赏价值。

然而，要在花卉种类识别任务中取得良好的性能，仍然面临一系列挑战。例如，花卉图像存在光照、角度、尺度等多样性，模型需要具备较强的鲁棒性。同时，花卉

种类的多样性和数量庞大也对模型的泛化能力提出了更高的要求。通过研究这些挑战，我们有望更好地理解深度学习在花卉种类识别中的应用，为未来研究提供更为准确和有效的方向。

总体而言，基于 ResNet 网络的花卉种类识别研究不仅在学术上具有重要的理论价值，也在实际应用中有着广泛的实用前景，对于推动植物学和计算机视觉领域的交叉研究具有积极的推动作用。

1.2 国内外研究现状

传统机器学习方法的主要研究内容如下：Nilsback 最早提出的花卉图像分割算法，用来分离前景和背景。之后 Chai 等人在 Nilsback 等的基础上进行改进，提出了一种联合过滤分割算法 BiCoS，Angelova 等人改进了之前的算法，他们先检测图像的 low-level 位置，接着用 Propagation 算法对原始的花卉图像进行全分割。而近年来，随着深度学习技术的发展，各种卷积神经网络都得到了极大的提升以及扩展，在卷积神经网络 CNN 为基础的深度学习方法中，以 GoogLeNet、ResNet、Batch Normalization 等网络模型为代表，并且取得了极大地进展。有学者采用 AlexNet 以及 VGG 等网络模型，用数据集训练好网络参数初始化网络层（最后一层全连接层除外），以较大的学习速率训练最后一层全连接层，以较小的学习速率微调瓶颈层；Hu 等人采用 GoogLeNet，先在数据集上训练网络，之后替换掉最后一层全连接层，再在数据集上微调网络；Wu 等人采用 ResNet50，先在数据集上进行网络的训练，之后以一个较小的学习速率再对所有网络层进行微调[1]。

在国外，花卉种类识别领域已经取得了显著的研究进展。研究者们普遍采用深度学习模型，特别是 ResNet 网络，来提高花卉识别的准确性和泛化能力。许多研究关注如何通过深度学习网络，尤其是 ResNet，在大规模花卉数据集上进行训练，以实现对数万种花卉的高效识别。这些研究通常包括对网络结构的调整和数据增强策略的优化。部分研究致力于解决小样本花卉数据集的问题，采用迁移学习和微调等技术，使得模型更好地适应具有限样本的特定花卉识别任务。

在国内，花卉种类识别的研究也逐渐兴起，研究者们开始关注如何利用深度学习技术提高花卉识别的精度。一些国内研究聚焦于优化 ResNet 网络结构，以适应中国特有的花卉种类和环境。这包括对网络深度、通道数等超参数的调整，以提高模型在本土花卉数据上的性能。国内的研究者还在探索如何将花卉识别技术应用于智能农业领域，例如自动采花系统的设计和实现。这方面的研究有望为农业生产提供更高效、

智能的解决方案[2]。

总体而言，基于 ResNet 网络的花卉种类识别研究在国内外都取得了显著的进展，研究者们持续致力于优化算法、拓展应用场景，为推动植物学和计算机视觉的交叉研究做出贡献。未来随着技术的不断创新，花卉种类识别技术有望在更多领域展现其实用性和智能化水平。

1.3 研究内容和结构

本研究的结构主要分为以下几个部分：

一、花卉数据集的构建：本研究首先致力于构建一个包含多个花卉种类的数据集，确保数据集的多样性和代表性。通过采集来自不同地理位置、气候条件和季节的花卉图像，以及考虑不同拍摄角度和光照条件，确保模型能够在真实世界的多样性中进行训练。

二、ResNet 网络的深入研究：在理论层面，研究将对 ResNet 网络的结构进行深入分析。探讨其残差学习机制，理解在深度网络中如何有效地传播梯度，解决梯度消失和梯度爆炸问题。通过对 ResNet 不同深度和复杂度的变种进行比较，选择最适合花卉种类识别任务的网络结构。

三、模型训练与优化：使用构建的花卉数据集，对选定的 ResNet 网络进行训练。通过使用深度学习 PyTorch 框架以及 GPU 加速，提高训练效率。在训练过程中，采用常见的数据增强技术，如旋转、缩放和翻转，以增加模型的泛化能力。优化过程中，调整学习率、使用合适的正则化方法，并考虑迁移学习，从预训练的模型中初始化参数，以提高模型收敛速度。

四、实验结果与分析：对训练好的模型进行系统的实验评估。通过与传统方法以及其他深度学习架构进行比较，分析 ResNet 网络在花卉种类识别上的性能表现。通过混淆矩阵、准确率、召回率等指标，全面评估模型在不同花卉类别上的识别效果。同时，对模型的鲁棒性和泛化能力进行测试，验证其在不同场景和条件下的适用性。

五、创新点的验证：验证研究中提出的创新点，包括深度学习技术的应用、多样性数据集的构建和模型优化策略。通过对比实验结果，证明 ResNet 网络在花卉种类识别任务上相较于传统方法和其他深度学习网络的优越性，并进一步解释其优越性的原因。

六、结果讨论与未来展望：在实验结果和分析的基础上，深入讨论研究的发现，指出模型的局限性和改进空间。同时，对未来研究方向提出建议，如探索更先进的深

度学习架构、进一步改进数据集的多样性，以及将模型应用于实际场景中的挑战和机会。

第 2 章 相关理论及方法介绍

本章主要对一些卷积神经网络模型进行介绍并对迁移学习等在研究中使用到的方法进行详细描述。

2.1 卷积神经网络

卷积神经网络（CNN）是一类专门设计用于处理具有网格结构数据的深度学习模型，尤其在计算机视觉领域中取得了显著的成功。CNN 的设计灵感主要来自于生物学中对动物视觉系统的理解，其中感受野和权值共享等概念在 CNN 中得到了应用[3]。

以下是 CNN 的一些关键概念和组成部分：

输入层是卷积神经网络各层中的底层，其功能是处理输入数据，如去除平均值，减少数据中的过度偏差，并改善训练效果。

卷积层（Convolutional Layer）： CNN 的核心是卷积层，通过卷积操作可以有效地提取输入数据的局部特征。卷积操作使用一个称为卷积核（或过滤器）的小矩阵，对输入数据进行逐元素相乘并求和的操作。通过在整个输入上滑动卷积核，可以得到输出特征图。卷积操作具有参数共享的特性，使得网络可以学习到局部特征的通用表示。其原理如下图所示。

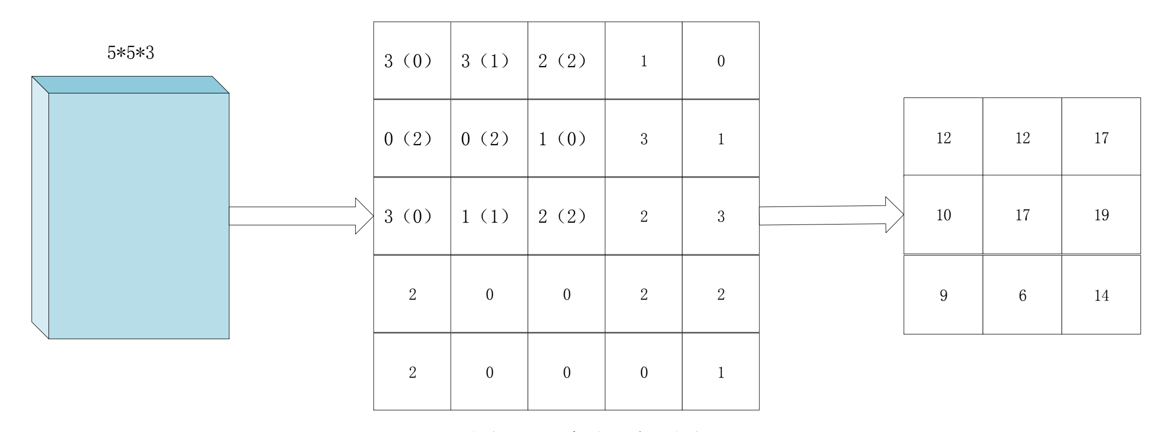


图 2.1 卷积原理图

池化层（Pooling Layer）：池化层用于降低卷积层输出的空间维度，减少计算量和参数数量，同时保留重要的特征。最大池化是常见的池化操作，它在每个区域内选择最大的值作为输出，从而保留区域内最显著的特征。平均池化是另一种常见的池化操作，采用区域内数值的平均值作为输出。最大池化原理如下图所示。

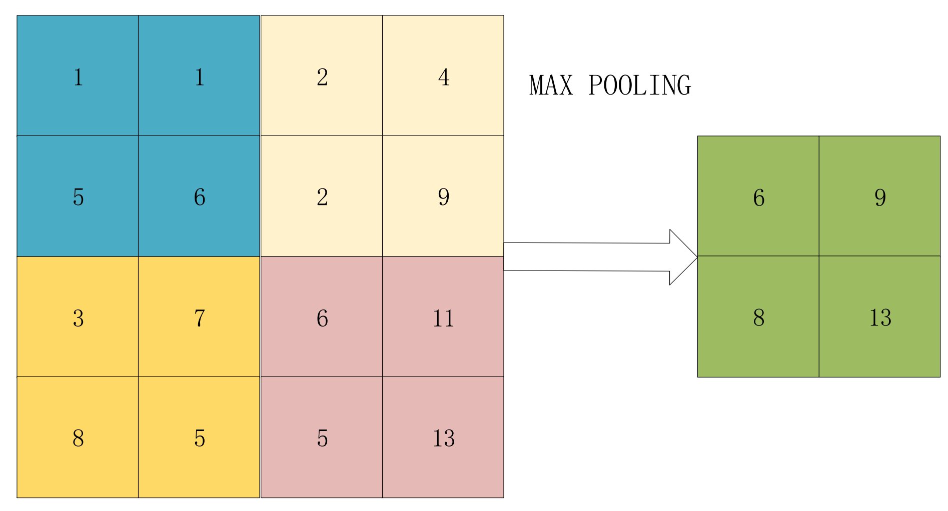


图 2.2 最大池化原理图

激活函数（Activation Function）：在卷积神经网络中，通常在卷积层和全连接层后添加激活函数，以引入非线性变换。常用的激活函数包括 ReLU（Rectified Linear Unit）、Sigmoid 和 Tanh。ReLU 是最常用的激活函数，它将负值置零，保留正值，有助于网络学习更复杂的表示。其函数图如下图所示。

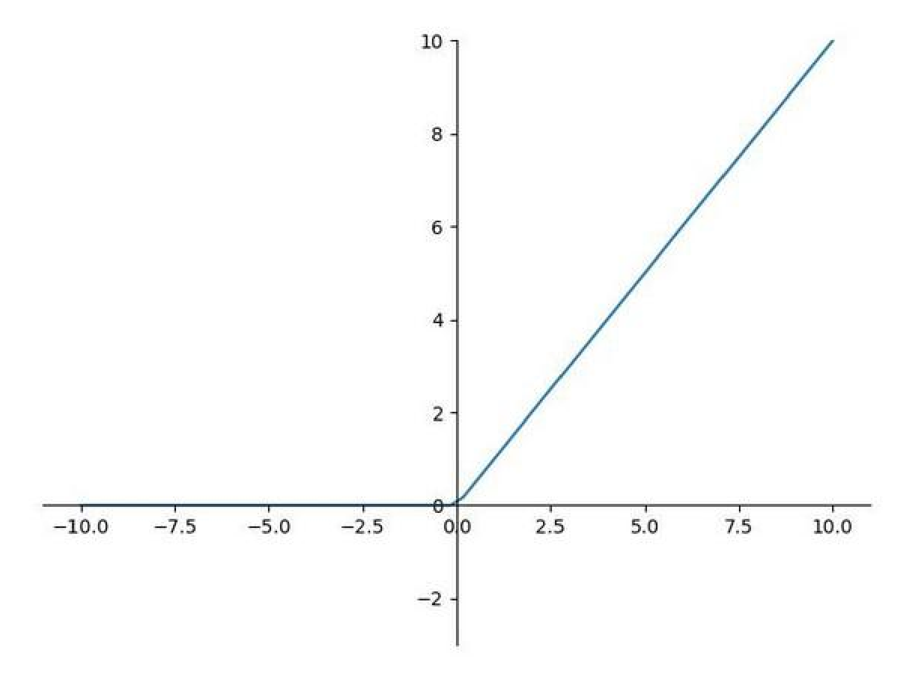


图 2.3 ReLU 函数图

全连接层（Fully Connected Layer）：全连接层通常接在卷积层和池化层之后，用于整合高级抽象特征并生成最终输出。全连接层中的每个神经元与前一层的所有神经元相连接，形成全连接的权重矩阵。

CNN 的设计使得它在图像识别、目标检测、语音处理等领域取得了广泛的成功，成为深度学习中的重要工具。其结构图如下图所示。

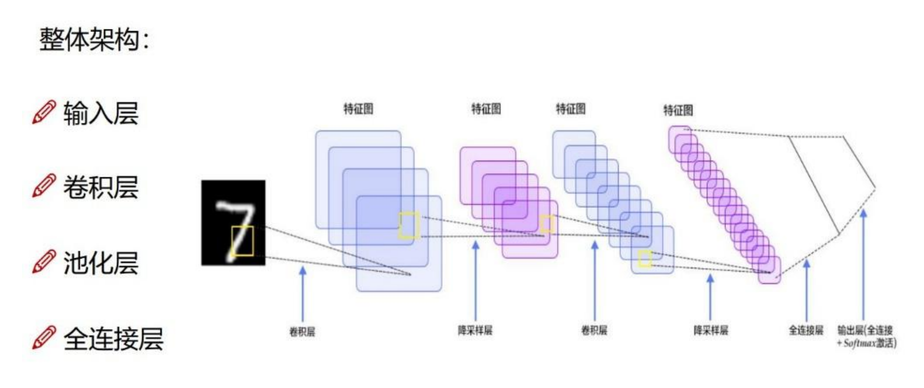


图 2.4 CNN 网络结构图

2.2 ResNet 网络

ResNet（Residual Network）是由 Microsoft Research 提出的一种深度学习网络结构，于 2015 年在 ILSVRC（ImageNet Large Scale Visual Recognition Challenge）比赛中取得了优胜。其主要贡献在于引入了残差学习的概念，通过残差块（Residual Block）的设计，使得网络可以更轻松地训练和优化超深的层次，解决了传统深度神经网络中的梯度消失和梯度爆炸等问题[4]。

以下是 ResNet 网络的主要特点和组成部分：

残差块（Residual Block）：ResNet 网络的核心是残差块，它包含了一个跳跃连接（shortcut connection）和两个普通的卷积层。跳跃连接允许输入直接绕过一个或多个层，将未经变换的输入加到输出上，这种结构允许网络学习到残差（或误差），从而更容易地训练深层网络。

层次结构（Layer Stacking）：ResNet 通过堆叠多个残差块来形成深层网络。这种层次结构使得网络可以根据任务的复杂性选择合适的深度，而不会受到梯度消失的影响。ResNet 的深度通常从几十层到数百层不等，相比于传统的深度网络，具有更强的表达能力。

全局平均池化（Global Average Pooling）：ResNet 中不使用传统的全连接层，而是采用全局平均池化层。这一层将每个特征图的所有元素取平均值，得到一个固定大小的输出，用于连接到最终的分类层。这样的设计减少了参数数量，有助于降低过拟合的风险。

超深的网络结构：ResNet 引入的残差学习机制使得网络可以轻松地超越百层，甚至千层。这种超深的网络结构有助于提取更高层次、更抽象的特征，从而提高模型的表达能力。

预训练与迁移学习：ResNet 网络通常通过在大规模图像数据集上进行预训练，例

如 ImageNet 数据集，以学习通用的特征表示。在训练目标任务时，可以采用迁移学习的方式，将预训练的 ResNet 网络的权重作为初始值，有助于提高模型的收敛速度和性能。

ResNet 的提出极大地推动了深度学习领域的发展，成为解决深度神经网络难以训练的问题的重要里程碑之一。它的成功启发了许多后续深度学习网络的设计，并在计算机视觉领域取得了许多重要的成果。其具体结构如下图所示[5]。

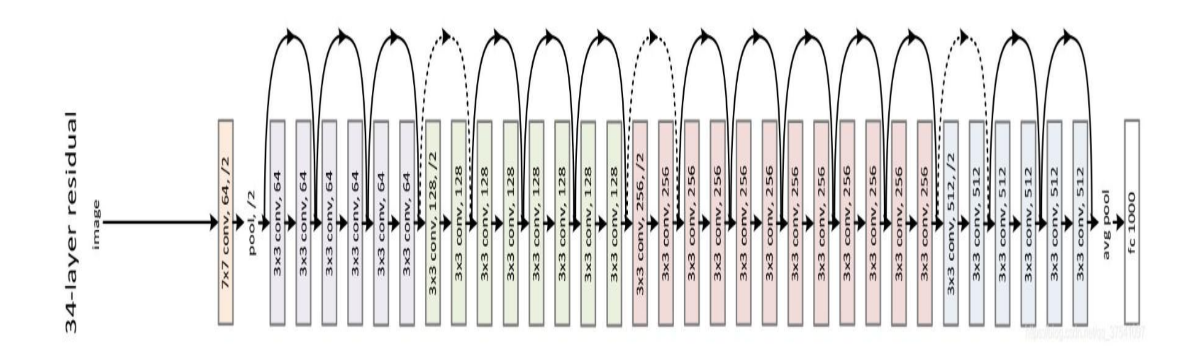


图 2.4 34 层 ResNet 网络结构图

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		[3×3, 64; 3×3, 64]×2	[3×3, 64; 3×3, 64]×3	[1×1, 64; 3×3, 64; 1×1, 256]×3	[1×1, 64; 3×3, 64; 1×1, 256]×3	[1×1, 64; 3×3, 64; 1×1, 256]×3
conv3_x	28×28	[3×3, 128; 3×3, 128]×2	[3×3, 128; 3×3, 128]×4	[1×1, 128; 3×3, 128; 1×1, 512]×4	[1×1, 128; 3×3, 128; 1×1, 512]×4	[1×1, 128; 3×3, 128; 1×1, 512]×8
conv4_x	14×14	[3×3, 256; 3×3, 256]×2	[3×3, 256; 3×3, 256]×6	[1×1, 256; 3×3, 256; 1×1, 1024]×6	[1×1, 256; 3×3, 256; 1×1, 1024]×23	[1×1, 256; 3×3, 256; 1×1, 1024]×36
conv5_x	7×7	[3×3, 512; 3×3, 512]×2	[3×3, 512; 3×3, 512]×3	[1×1, 512; 3×3, 512; 1×1, 2048]×3	[1×1, 512; 3×3, 512; 1×1, 2048]×3	[1×1, 512; 3×3, 512; 1×1, 2048]×3
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

图 2.5 各层对应结构表

在本次的实验中使用的 ResNet 网络模型为 152 层的网络模型。

2.3 其他网络模型

本实验为了更好的展示 ResNet 网络在进行识别方面的优越性，使用了 AlexNet 网络、VGG 网络等模型，其中 AlexNet 网络和 VGG 网络结构如下图所示。

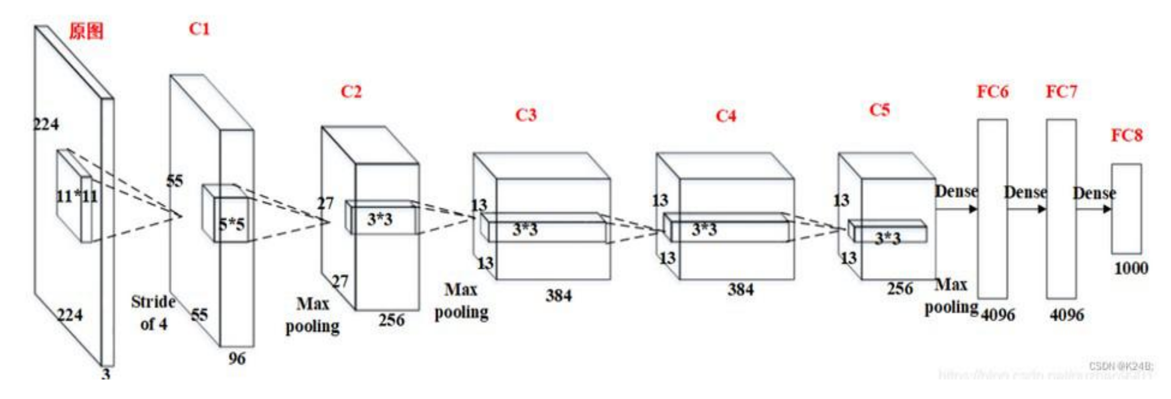


图 2.6 AlexNet 网络结构图

AlexNet 是一个相对较深的卷积神经网络，具有 8 个学习层（5 个卷积层和 3 个全连接层）。AlexNet 使用了较大尺寸的卷积核，如 11x11、5x5 和 3x3。这有助于捕捉不同尺度的特征。AlexNet 使用 ReLU（Rectified Linear Unit）作为激活函数，这有助于引入非线性，加速训练过程。AlexNet 引入了局部响应归一化层，这有助于增强模型的泛化能力。然而，在后续的研究中，LRN 并没有被广泛采用，而是被其他正则化技术如批标准化（Batch Normalization）所取代。在卷积层之后，AlexNet 使用最大池化层进行下采样，以减小特征图的尺寸。AlexNet 包含三个全连接层，其中最后一个全连接层输出 1000 个类别的概率分布，用于 ImageNet 的分类任务。在全连接层中，AlexNet 采用了 Dropout 层，以减少过拟合。AlexNet 在训练时使用了大量的数据增强技术，如翻转、随机裁剪等，以增加训练数据的多样性，提高模型的泛化能力。

总体上，AlexNet 的成功标志着深度学习在计算机视觉领域的崛起，并为后续更深、更复杂的神经网络奠定了基础。由于计算资源和数据集的限制，AlexNet 的设计考虑了如何充分利用 GPU 进行训练，这也促使了深度学习硬件和软件的发展。

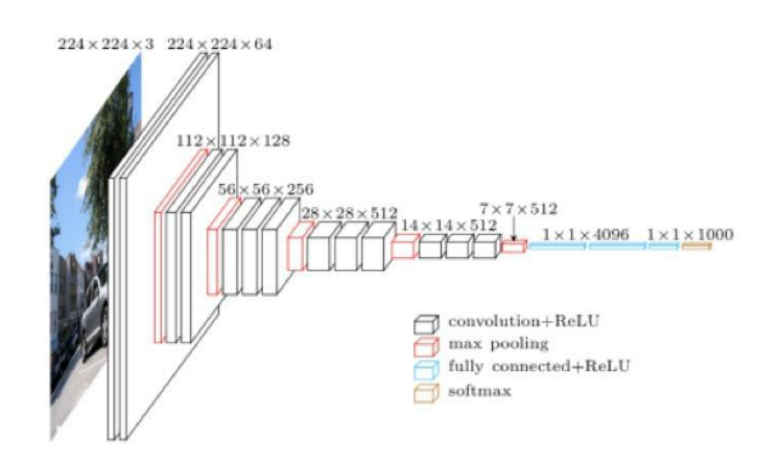


图 2.7 VGG 网络结构图

VGG（Visual Geometry Group）是一种卷积神经网络（CNN）架构，由牛津大学

的 Visual Geometry Group 团队提出。该架构首次在 2014 年的 ImageNet 大规模视觉识别竞赛（ILSVRC）中取得了显著的成功，VGG16 和 VGG19 是最为知名的两个版本。

VGG 网络相对较深，其核心思想是通过使用小尺寸的卷积核和较深的网络来提高性能。VGG16 包含 16 个卷积层和全连接层，而 VGG19 则包含 19 个。VGG 网络中的卷积层均使用 3x3 的卷积核，并采用步幅为 1 和零填充。这种设计有助于保留空间信息，同时减少参数数量。在卷积层之后，VGG 使用最大池化层进行下采样。最大池化用于减小特征图的尺寸，从而降低计算负担和提高网络的感受野。VGG 网络在卷积层之后包含全连接层，用于整合高层次的语义信息。这些全连接层在最后几层的特征图上操作，以生成最终的分类输出。VGG 网络中常使用 ReLU（Rectified Linear Unit）作为激活函数，以引入非线性性。在全连接层中，VGG 网络使用了 Dropout 层，有助于防止过拟合。最后的全连接层用于执行分类任务，通常采用 softmax 激活函数输出类别概率。由于 VGG 相对简单而直观的架构，它成为了深度学习入门和理解的良好范例。

2.4 迁移学习

迁移学习（Transfer Learning）是一种机器学习方法，它利用在一个任务上学到的知识，来改善在新任务上的学习性能。在深度学习领域，迁移学习被广泛应用，尤其在数据集较小、计算资源有限的情况下，能够有效提高模型的泛化能力[5]。

迁移学习旨在通过将在一个相关任务上学到的知识应用于新任务，从而改善新任务的学习性能。迁移学习的关键假设是，不同的任务之间存在一些共享的知识，这些知识可以帮助提高模型在新任务上的性能。迁移学习的常见做法之一是使用在大规模数据集上预训练的模型，如 ImageNet 上预训练的卷积神经网络（CNN）或 BERT 等自然语言处理任务上预训练的模型。这些预训练模型已经学到了丰富的特征表示，可以作为通用的特征提取器。在迁移学习中，常常使用层的冻结和解冻策略。冻结一些层的参数意味着在训练过程中保持它们不变，只训练部分层的参数，通常是新添加的全连接层。随着训练的进行，逐渐解冻之前冻结的层，使得整个模型逐渐适应新任务。在迁移学习中，需要根据目标任务的性质来评估模型性能。在微调过程中，可能需要进行超参数的调整，例如学习率、训练批次大小等，以便更好地适应新任务[6]。

迁移学习在实际应用中得到了广泛的应用，特别是在计算机视觉、自然语言处理和语音识别等领域。它不仅可以加速模型的训练过程，同时也能够提高模型对新任务的泛化能力。

第 3 章 数据处理

本章对实验中的图像数据库的构建，图像的预处理以及调用等操作进行简要介绍，并展示了进行与处理后的图像。

3.1 图像数据库建立

本实验共收集了 7000 余张图，品种达到 100 多种。在构建图像识别系统时，图像数据收集一般不是在环境中搜索对象，就是事先收集对象实例。数据集的质量很大程度上取决于收集和注释的性能。实验中使用的数据集分为训练集(train)和验证集(valid)两部分，这两种数据集共同存放于“flower_data”文件夹下，训练集和验证集每个文件下都包含一定数量的子文件夹，并且这些子文件夹按照顺序进行命名，每个子文件夹下存放有同一品种的花卉图片若干张。

在建立完训练和验证所需要的数据集后需要建立一个 name 文本，格式为 json 文件，文本内容以字典的形式进行存储，分别对应子文件夹下的每种花卉的名字。建立该文件的目的是在进行输出是对应花卉的名称以便于观察是否准确。具体如下图所示。

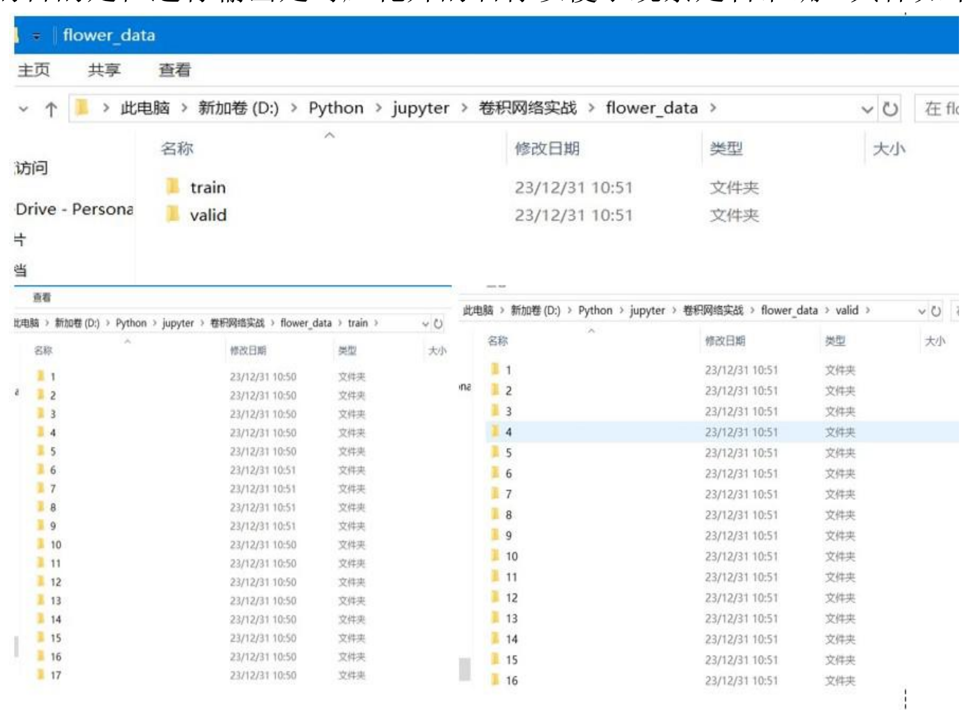


图 3. 1 数据集文件

cat to name.json - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

{"21": "fire lily", "3": "canterbury bells", "45": "bolero deep blue", "1": "pink primrose", "34": "mexican aster", "27": "prince of wales feathe
permum", "70": "tree poppy", "85": "desert-rose", "99": "bromelia", "87": "magnolia", "5": "english marigold", "92": "bee balm", "28": "ste
"88": "cyclamen", "94": "foxglove", "81": "frangipani", "74": "rose", "89": "watercress", "73": "water lily", "46": "wallflower", "77": "passion f

图 3.2 文件夹对应花卉名称

3.2 数据的预处理

由于本次所用于训练的数据集并不多，为了使训练达到最大化，需要对数据进行预处理操作。数据的预处理主要有中心裁剪，图像翻转等，目的是使数据具有多样性，以保证训练效果。同时数据的预处理可以减少训练神经网络的计算时间，提高读取数据的效率。在本次的实验中，对数据进行预处理、数据的增强以及预处理等操作主要是通过 TorchVision 中的 Transform 模块进行实现的。在本次的实验中，图像预处理采用-45 到 45 度之间随机旋转，长度为 224 的中心裁剪，并且选择概率为 0.5 进行随机水平翻转、垂直翻转以及概率为 0.025 转化成灰度率，在训练时使用的 batch_size 为 8。具体数据处理结构图下图所示。

```
5]: image_datasets

5]: {'train': Dataset ImageFolder
         Number of datapoints: 6553
         Root location: ./flower_data/train
         StandardTransform
     Transform: Compose(
                    RandomRotation(degrees=[-45.0, 45.0], interpolation=nearest, expand=False, fill=0)
                    CenterCrop(size=(224, 224))
                    RandomHorizontalFlip(p=0.5)
                    RandomVerticalFlip(p=0.5)
                    ColorJitter(brightness=(0.8, 1.2), contrast=(0.9, 1.1), saturation=(0.9, 1.1), hue=(-0.1, 0.1))
                    RandomGrayscale(p=0.025)
                    ToTensor()
                    Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])
                ),
     'valid': Dataset ImageFolder
         Number of datapoints: 818
         Root location: ./flower_data/valid
         StandardTransform
     Transform: Compose(
                    Resize(size=256, interpolation=bilinear, max_size=None, antialias=warn)
                    CenterCrop(size=(224, 224))
                    ToTensor()
                    Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224, 0.225])
                )}
```

图 3.3 训练集和验证集处理

3.3 数据展示

在进行数据预处理完操作后，可以观察到处理后的图像如下。

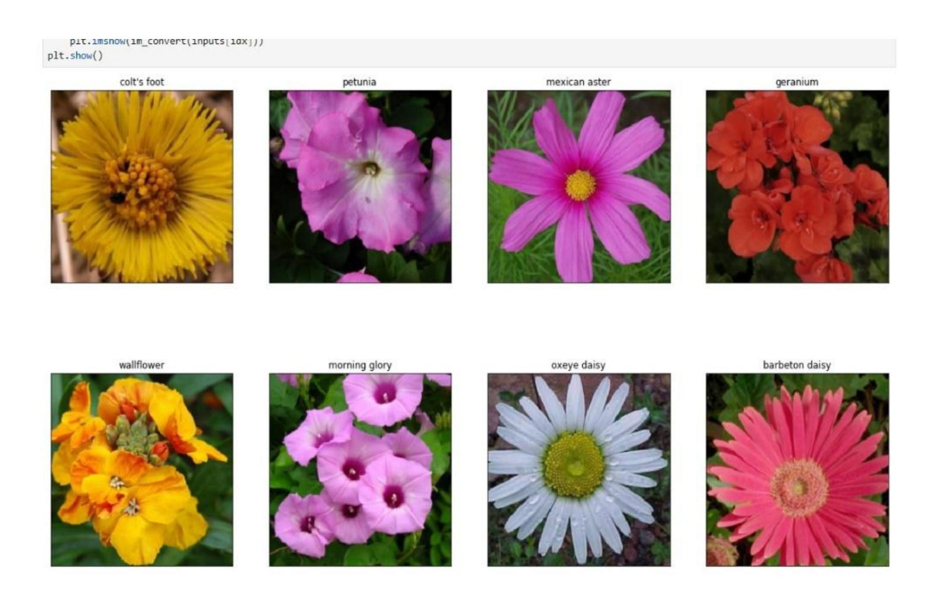


图 3.4 处理后图像展示

通过观察可以看出，在经过预处理之后图像全部变成了大小相同的格式，且花卉与 name 文件中的花卉名称相对应。

第 4 章 网络搭建及模型训练

本文主要介绍神经网络模型的选取搭建以及对模型的训练。

4.1 网络模型选取及搭建

本研究中所有的模型均是以 PyTorch 框架进行构建的，主要 3 种网络模型，具体的模型选择可以通过对“model_name”的定义进行实现，如下图所示。

```
model_name = 'resnet'  #可选的比较多 ['resnet', 'alexnet', 'vgg
#是否用人家训练好的特征来做
feature_extract = True
```

图 4. 1 网络模型选择

在神经网络模型训练过程中，相比于 CPU 而言，GPU 可以极大提升训练效率，因此在进行模型搭建前需要判断设备是否具备 GPU 运算能力。在开始时会加载 models 中提供的模型，并且直接用训练的好权重当做初始化参数。在以上过程完成之后就可以进行网络模型的构建，由于 ResNet 网络、AlexNet 网络以及 VGG 网络模型均有训练好的，所以只需通过迁移学习将其下载下来并按照实验需要进行修改即可。以 ResNet 网络模型为例子，这里对其需要进行训练的层进行了修改，连接层修改为了 102。

4.2 参数设置

4.2.1 输出层与梯度设置

为了防止每一层计算过后梯度会进行更新，这里将每一层的“param.requires_grad”参数设置为 False，这样在进行反向传播的时候会停止计算梯度，使其保持不变。当将“param.requires_grad”的参数值改为 False 时，运行可以发现不再显示每层的参数，表示不再进行更新。

4.2.2 优化器与学习率衰减

在本次实验中，网络模型使用的优化器为 Adam 优化器，学习率经过查阅资料以及多次试验设置为 0.001，学习的衰减率为 0.1，具体方式如下图所示。

```
[43]: # 优化器设置
optimizer_ft = optim.Adam(params_to_update, lr=1e-2)
scheduler = optim.lr_scheduler.StepLR(optimizer_ft, step_size=7, gamma=0.1)#学习率每7个epoch衰减成原来的1/10
#最后一层已经LogSoftmax()了，所以不能nn.CrossEntropyLoss()来计算了，nn.CrossEntropyLoss()相当于LogSoftmax()和nn.N
criterion = nn.NLLLoss()
```

图 4.2 优化器及衰减率设置

4.3 模型训练

在每个训练周期中，设置“running_loss”和“running_corrects”的值为零，之后进行训练集和测试集的加载，在每个周期中，只有在训练的时候才会计算和更新梯度。之后会计算损失并得到最好的那次模型，训练完成之后将最好的一次当做模型训练的结果。

4.4 结果展示

在本次实验中，由于所用设备落后以及数据资料等方面的原因，经过多次验证选择训练周期为 45，经过不同模型训练后，使用一个 batch 中的数据进行花卉识别效果如下图所示。















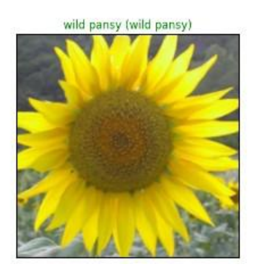


图 4.3 ResNet 网络训练结果预测

















图 4.4 VGG 网络训练结果测试















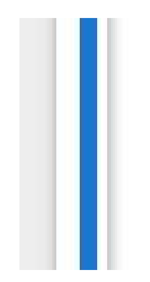


图 4.5 AlexNet 网络训练结果测试

其不同模型经过相同的训练周期后最终识别的准确率如下图所示。

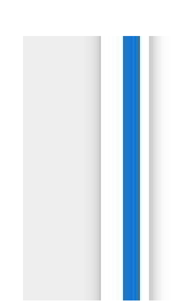
```
----------
train Loss: 3.4059 Acc: 0.8086
valid Loss: 8.7821 Acc: 0.6528
Optimizer learning rate : 0.0010000
```

图 4.6 AlexNet 网络



```
----------
train Loss: 2.9266 Acc: 0.7981
valid Loss: 9.3593 Acc: 0.6455
Optimizer learning rate : 0.0010000
```

图 4.7 VGG 网络



```
----------
train Loss: 2.6867 Acc: 0.8265
valid Loss: 8.2074 Acc: 0.6932
Optimizer learning rate : 0.0010000
```

图 4.8 ResNet 网络

不同模型下得到的准确度如下表所示。

表 4.1 模型训练测试准确度

模型	训练集	测试集
ResNet152	83%	74%
AlexNet	80%	65%
VGG16	80%	64%

通过训练后测试集最终的准确率可以看出，相比于其他的神经网络，ResNet 网络模型的准确率更高，更适合用于处理花卉识别。

由于数据集以及设备原因导致训练周期有限，所以各模型的识别准确率并未达到最大，但通过每一周期训练得到的数据趋势来看，ResNet 网络模型是目前比较适合进行花卉识别训练的模型。

结 论

通过观察实验得出的数据可以看出，基于 ResNet 网络的花卉种类识别研究相比于其他网络结构主要有以下几个优势。

提高了识别准确性：实验结果表明通过使用 ResNet 网络在花卉种类识别任务上取得了更高的识别准确性，ResNet 网络的深度和残差学习机制使其在学习复杂的花卉特征时更为有效，相较于传统方法或其他深度学习网络，表现更为优越。

网络深度对性能的影响：通过对 ResNet 网络的深度进行实验，可能得出结论指出，在一定范围内增加网络深度有助于提高花卉种类识别的性能。然而，过深的网络可能导致过拟合或训练困难，因此需要在深度和性能之间找到合适的平衡点。

迁移学习的有效性：如果在研究中采用了迁移学习策略，结论可能会强调其在花卉种类识别任务上的有效性。预训练的 ResNet 网络在大规模图像数据集上学到了通用的特征表示，为花卉识别任务提供了有力的先验知识，从而加速模型的收敛并提高泛化性能。

参考文献

[1] 付雪婷,王新鑫,杨凡凡等.基于卷积神经网络的植物品种识别研究[J].南方农机,2023,54(17):65-69.

[2] 刘景贺,任彦彪,薛岩等.基于 ResNet152 的花卉识别研究[J].电脑知识与技术,2023,19(15):15-17.

[3] 廖明霜,罗远远.基于 ResNet 对花朵分类研究[J].农业与技术,2023,43(02):65-68.

[4] 王雪琰,张冲,张立.卷积神经网络下的相似月季识别[J].安徽农业大学学报,2021,48(03):504-510.

[5] 张梦雨.基于 ResNet 和注意力机制的花卉识别[J].计算机与现代化,2021,(04):61-67.

[6] 关胤.基于残差网络迁移学习的花卉识别系统[J].计算机工程与应用,2019,55(01):174-179.

[7] 刘凡,陈锐.基于改进 ResNet50 在高岭土矿石图像分类的应用[J/OL].机电工程技术,1-7[2024-01-01]

[8] 陈卫国,莫胜撼.基于迁移学习和逻辑回归模型的花卉分类研究[J].南方农机,2024,55(01):139-143+151.

[9] 付清华.基于迁移学习的卷积神经网络花卉识别研究[J].科学技术创新,2023,(18):112-115.

[10] 赵洋,梁迎春,许军等.改进 ResNet18 网络模型的花卉识别[J].计算机技术与发展,2022,32(07):167-172.

[11] 王芳,郑圣勇.基于改进的 VGG-16 模型的花卉识别小程序设计[J].信息与电脑(理论版),2022,34(11):157-159.

[12] 张旭.基于卷积神经网络的花卉识别算法研究[D].内蒙古大学,2020.

[13] 秦敏.基于深度学习的花卉图像分类识别模型研究[D].广西师范大学,2020.

[14] 宋子龙.基于卷积神经网络的花卉种类识别系统[J].计算机产品与流通，2019(12)：91+131.

[15]刘晓曈.基于深度学习的分类预测方法研究及应用[D].南京:东南大学，2017

致 谢

在完成本研究论文的过程中，我们深感荣幸，并要向许多人表达我们最诚挚的感谢之情，是你们的支持和帮助使得我们的研究得以成功完成。

首先，我们要衷心感谢我的导师。导师不仅在学术研究方面有着卓越的造诣，更在整个研究过程中给予了我们的极大关心和支持。他的悉心指导、深刻的见解和对学术研究的热情激发了我们不断前行的动力。导师在每个阶段都给予了宝贵的建议，使得我们的研究能够更加深入和完善。

与此同时，我要感谢实验室的所有成员和各位同学，他们为研究提供了良好的学术氛围和团队协作精神。大家在实验室共同努力，相互交流，共同克服了研究中的难题，使得整个研究过程变得更加丰富和有趣。在实验室的学术氛围中，我学到了很多宝贵的知识和技能。此外，我要特别感谢数据集提供者，他们的数据为我的实验提供了基础。数据集的质量和多样性为我的模型训练和评估提供了充足的支持。在研究过程中，我从数据集中获取了丰富的图像样本，这对于我的研究起到了至关重要的作用。

最后，我们要向开源社区表示由衷的感激。ResNet 网络作为一个开源项目，为我的花卉种类识别任务提供了强大的网络结构。感谢 ResNet 网络的创建者和贡献者，你们的工作为我们的研究提供了可靠的技术基础，加速了我们的研究进程。

在这个研究的旅程中，我们遇到了挑战，也收获了成功。感谢所有在我的研究中给予帮助和支持的人们，是你们的付出使得我能够顺利地完成了这项研究工作。